Our words, our story: a textual analysis of articles published in the *Bulletin of the Medical Library Association/Journal of the Medical Library Association* from 1961 to 2010*

Mark E. Funk, MA, AHIP, FMLA

See end of article for author's affiliation.

DOI: http://dx.doi.org/10.3163/1536-5050.101.1.003

Purpose: This lecture explores changes in the medical library profession over the last fifty years, as revealed by individual word usage in a body of literature.

Methods: I downloaded articles published in the *Bulletin of the Medical Library Association* and *Journal of the Medical Library Association* between 1961 and 2000 to create an electronic corpus and tracked annual frequency of individual word usage. I used frequency sparklines of words, matching one of four archetypal shapes (level, rise, fall, and rise-and-fall) to identify significant words.

Results: Most significant words fell into the categories of environment, management, technology, and research. Based on word usage changes, the following trends are revealed: Compared to 1961, today's medical librarians are more concerned with digital information, not physical packages. We prefer information to be evidence-based. We focus more on health than medicine. We are reaching out to new constituents, sometimes leaving our building to do so. Teaching has become important for us. We run our libraries more like businesses, using constantly changing technology. We are publishing more research articles.

Conclusions: Although these words were chosen by individual authors to tell their particular stories, in the aggregate, our words reveal our story of change in our profession.

The Janet Doe Lecture on the history or philosophy of medical librarianship: I will warn you right now that you will get very little philosophy out of me today, for two reasons. First, my predecessor T. Scott Plutchak, AHIP, was a philosophy major, and I knew that a zoology major like myself could not compete with that. Second, ''philosophy'' in many Janet Doe lectures is actually a strong personal viewpoint of medical librarianship—what it is or what it should be. I do collection development, where things change so much and so rapidly that I have not had time to develop a strong viewpoint like many of my Doe predecessors. Like most of my collection development colleagues, I'm just trying to survive day to day. That tends to create a very pragmatic attitude. If I believe in anything strongly, it is that I believe I'll have another cookie.

Without philosophy, I am left with history. And here I will echo the complaint of many Doe lecturers by stating that I have a severe lack of historical research skills. I became painfully aware of this lack as I read previous Doe Lectures, such as David Kronick's 1980 lecture [1]. Kronick was a true scholar, with a doctorate in librarianship. We honor him to this day with the Medical Library Association's (MLA's) David A. Kronick Traveling Fellowship. In his Doe Lecture, Kronick quoted H. Curtis Wright's ''The Oral Antecedents of Greek Librarianship,'' Francis Bacon, and the fifteenth-century Abbot Johannes Trimethius. In contrast, later in this speech, I will quote the Talking Heads.

While I am totally unqualified for traditional historical research, that still leaves informal, or personal, history. Although I am old enough to be in my anecdotage, I just do not have many interesting stories to tell. And as Thomas Basler, FMLA, told us in his 2008 Doe Lecture, there are no more giants. While I met some of those giants, I did not know them, and I certainly do not have any stories to tell about them. I suppose I could tell stories about some of the taller than average individuals I have met in my career, but that does not sound very exciting.

BACKGROUND

With these limitations, I decided to try a different kind of historical research. I ran across an article in a 2011 issue of *Science* called ''Quantitative Analysis of Culture Using Millions of Digitized Books'' [2]. These researchers, with the assistance of Google, created a corpus of over 5 million digitized texts, representing about 4% of all books ever printed.†

By tracking the changing frequency of word usage over time, they could investigate cultural threads quantitatively. For example, they found a dramatic rise in frequency of the word ''slavery'' in the years leading up to the Civil War, with a sharp decrease in frequency immediately following the war. They also discovered a shorter rise in frequency during the Civil Rights Era of the late 1950s and 1960s. In another example, they found a huge gap between the relatively

* The Janet Doe Lecture on the history or philosophy of medical librarianship, presented at MLA '12, the 112th Annual Meeting of the Medical Library Association; Seattle, WA; May 21, 2012; T. Scott Plutchak, AHIP, the 2011 Janet Doe Lecturer, gave the introduction.

Supplemental Appendix A and Appendix B are available with the online version of this journal.

† The Google Ngram Viewer <http://books.google.com/ngrams/> came out of this project, where you can do a search on your own words.

high usage of the word ''men'' and the low usage of the word ''women'' from 1800 until the late 1960s, when the Feminist Movement began. After that, ''women'' was used more often in books than ''men.'' From these examples, one realizes that word usage does reflect real-life history.

I liked this approach. It did not require any philosophy, nor did it require historical knowledge of ancient Greek librarianship. I also thought that I could avoid the blind men and the elephant problem (making determinations based only on partial information). It would allow me to take an objective look at our history and not just reemphasize my own viewpoint.

All I needed was a large digitized corpus representing our profession. Where could I find one? I did not have to look very far. Estelle Brodman, the editor of the *Bulletin of the Medical Library Association (BMLA)* from 1949–1957, believed that the *BMLA* had several responsibilities. It should ''be an organ for reflecting the contemporary professional scene; it should put before its readers the newer advances in its field; and it should in some way present its profession in relation to other disciplines and other forces of society'' [3]. In other words, it should tell our story.

The *BMLA* and the *Journal of the Medical Library Association (JMLA)* sounded like the ideal bases for a digitized corpus that would accurately represent our profession. Could I use this quantitative technique to produce an interesting historical and cultural analysis of the changes in our profession? Could I use just our words to tell our story? Let us find out.

METHODOLOGY

How did I do this? I broke this down into four steps.

- Download *BMLA/JMLA* articles from PubMed Central into my computer.
- Copy and paste text from portable document format (PDF) files into Microsoft Word.
- Analyze word use frequency.
- Create Doe Lecture.

While all of the *BMLA/JMLA* issues are digitized in PubMed Central, I chose the years 1961 to 2010. Fifty years was a nice round number, and I think those decades represent a time of many changes in our profession. Change is good for a story. As an aside, I will point out that after I decided to start with 1961, upon opening that issue in PubMed Central, a frisson rushed through me when I realized the first article in that issue was written by Janet Doe herself.

I did not take everything that was published in those fifty years. I took articles and brief communications. I also included editorials and letters, but only if they did not concern MLA business. I wanted content about our *profession*, not our *association*. This meant I excluded certain things. I excluded book reviews and letters responding to book reviews. Also excluded were the proceedings of the annual meeting and reports from other meetings. Obituaries, association news, and president's pages were not included. For articles that were primarily bibliographies, I only included the prefatory text (e.g., all of the editions of the Brandon/Hill lists). Within articles, I skipped tables, appendixes (whether printed or online), footnotes, and references.

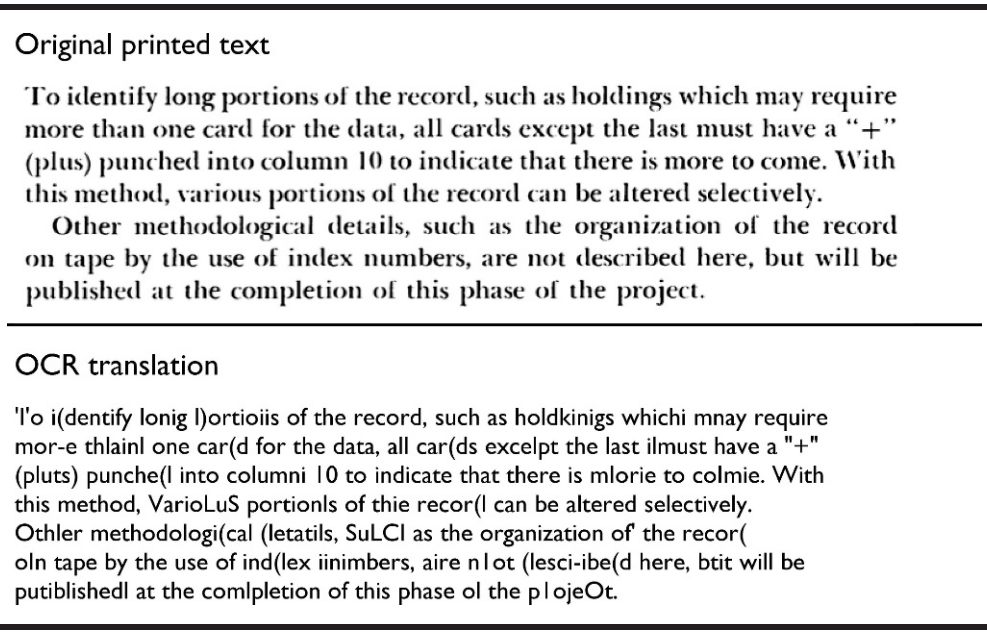
I have some caveats about this study. This is not a domain analysis. Neither is it an ontology. It is simply a frequency analysis of individual words. Even though many words can have multiple meanings, I felt that a few instances of different meanings would not matter much over fifty years, especially since many of these words were used thousands of time. This is a *very* brief report of some interesting patterns of word usage. There is much more that can be gleaned from this corpus, and I hope others will follow up, perhaps creating a domain analysis or an ontology. To that end, I am making the entire corpus available for anyone who wants to do further research. Text files and a comma delimited file organizing the word count by year are available in Appendix A (online only).

Copying and pasting the text from the PDFs into Word was not as straightforward as I had expected. The year 2000 was an important milestone in my methodology, since after that year the *BMLA* and *JMLA* articles in PubMed Central were PDF-native. They had always existed in digital form, which meant the text I copied was very accurate, requiring little correction by me. Before 2000, the articles were searchable PDFs, which sometimes had problems with the accuracy of the text, requiring manual correction. Why did this happen? The National Library of Medicine contracted to have all of the issues of the *BMLA* and *JMLA* scanned, using optical character recognition (OCR) software to produce searchable PDFs. Fortunately, most of the time the OCR software produced accurate text translations. However, when there were problems with the clarity of the original scanned copy, the OCR software produced wildly inaccurate text. Figure 1 is an example from 1963. In the original text seen on the top, the printing is uneven, with some thin lines and broken serifs. Although a human can read this text quite easily, the OCR software produced the translation seen on the bottom. Spellcheck could not fix this; it had to be manually corrected. This is from 1963, just the third year into my fifty-year project, and after running into several pages like this, I despaired that I would ever finish. Fortunately, the printing quality improved, which allowed the OCR software to produce more accurate translations.

Other problems with the OCR translations included some two-column pages being ''read'' by the software as a single column; and all em dashes—used to indicate a break in thought—being translated as hyphens, making for hundreds of very creatively hyphenated words. I wanted as accurate a representation of our words in the corpus as possible, which meant all of these problems had to be fixed manually. These led to what I called the ''Lost Summer of 2011.''

It took over 225 hours to create and edit the corpus. When finished, the corpus consisted of over 7.6 million words, taking up 18,648 single-spaced pages in Word (using the default Courier 10.5 font). I created a stop list of 943 words, and after applying it, ended up with 84,436 unique words, with the most frequent word used 65,174 times.

Figure 1
Original scanned text from 1963 *Bulletin of the Medical Library Association* article on top and the resulting optical character recognition translated text on bottom



With 84,436 unique words and 50 columns of years, my Excel spreadsheet had almost 4 and a quarter million data points. How could I narrow this down? I wanted to find trends in usage, and I realized that rarely used words would not show trends, so I decided to eliminate words that were used fewer than 50 times between 1961 and 2010. While this gave me a more realistic database of 6,893 words, even that number was large enough to make it difficult to decide which words to consider. I could have decided to just pick some words that I was interested in and follow their trends, but that only brings back the blind men and the elephant problem. I wanted the *corpus* to tell me what was important.

I decided that sparklines could help me make sense of this data. Edward Tufte, the guru of data and information presentation, introduced sparklines in his 2006 book, *Beautiful Evidence* [4]. Tufte defines sparklines as data-intense, design-simple, word-sized graphics. I used Excel to create a sparkline for each of the unique words. By looking for specific patterns in these sparklines, I could narrow down my search for significant words.

I focused on four archetypal patterns of sparklines for my analysis:

- Level: which can be interpreted as core, unchanging interest or activity in a word or topic
- Rise: an increasing interest or activity
- Fall: a decreasing interest or activity
- Rise-and-fall: a period of increasing interest or activity, followed by decreasing interest or activity

RESULTS

After I examined the sparklines that fit the archetypes, I realized that most of the words fit into four main categories, which correspond to major plot points in our story. Environment is where we live, in both the physical and conceptual sense. To understand and improve our environment, we make use of three major tools: management, technology, and research.

Environment

Medical librarianship exists in multiple environments. Think about your library. It is part of the general world of librarianship. It also exists within your institutional environment, and yet it is also in the much larger biomedical environment. As these environments have changed over the last fifty years, so has our story, as expressed in our words.

Socrates said an unexamined life was not worth living. Our lives must be worth living, since we examine ourselves quite regularly. **Librarian** and **librarians** are obvious words for talking about ourselves, and combining the counts of these two words reveals a slight but noticeable increase in their usage over the past twenty-five years.‡

We do seem to be a bit obsessed about our dwelling. The number one word we used over those fifty years was **library,** with the plural **libraries** coming in at number five. Combining the counts for both words reveals a rise in usage through the 1960s and 1970s. This represented a time of construction of new library buildings, and we liked to tell each other about our new houses. Our construction phase is over, and we have now returned to the same frequency of the early 1960s.

‡ Boldface words indicate the words that I tracked usage of in the corpus. Graphs for these words can be seen in Appendix B, online only.

When the graphs for librarian and library are overlaid (Figure 2), the trend lines show the gap between buildings and people is narrowing. Last year, Scott Plutchak reminded us that buildings do not do things; people do things [5]. This graph indicates that in another decade or two, we might be talking more about people than buildings. Let us move on to specific library tasks.

On the technical services side of things, we do not write as much about the **catalog** or **cataloging** as we used to. But that decrease makes sense, because we also do not write nearly as much about **books** anymore. This is a dramatic decrease in a topic that most people typically associate with us. Those stereotypes are in dire need of updating.

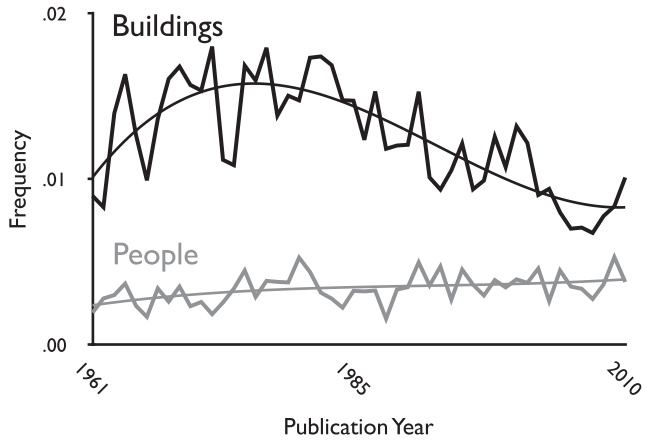
On the other hand, we are much more interested in talking about **journals.** In the overlay in Figure 3, you can see that the trend lines for books and journals crossed in the early 1970s and have diverged ever since.

Our writing about **acquisitions** has decreased quite a bit since the late 1970s, but this term is being replaced by our increasing interest in **licensing** our resources, particularly as more of them became electronic in the 1990s.

In public services, **reference** has been fairly steady over the decades, although there have been obvious changes in reference activities. There is less of an interest in **bibliographies** and things **bibliographic**, which have been more than replaced by writing about **searching.**

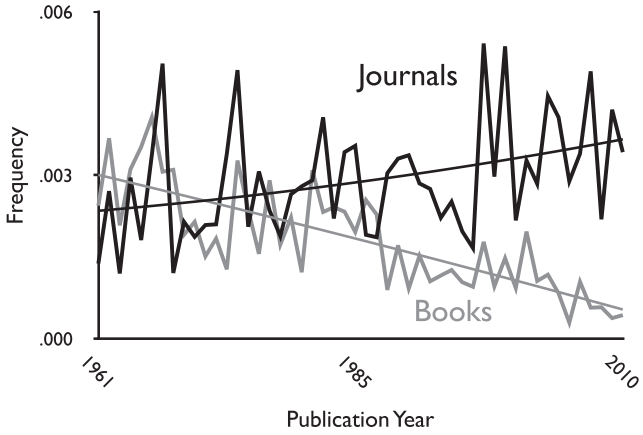
Much of our searching is in **MEDLINE.** Since MEDLINE became available in the 1970s, interest in it and its related products, **PubMed** and **MedlinePlus,** has seen a steady climb. With the amount of literature that can be retrieved in a search, we are becoming more concerned with helping our users not only to **locate** the information they need, but in recent years

Figure 2
Change in usage frequency* between words for buildings (library and libraries) and words for people (librarian and librarians)



* The number of times the words appeared in a given year divided by the total number of words for that year.

Figure 3
Change in usage frequency* between the words journals and books (singular and plural)



* The number of times the words appeared in a given year divided by the total number of words for that year.

helping them to **filter** the massive amounts of literature that are often retrieved.

A growing trend is to talk about the **resources** that a library has to offer, both to our users and to other libraries. But we are primarily interested in the **information** that is contained inside these resources. Information was the second most used word in the corpus, second only to library. However, just knowing where to find this information is not enough anymore. Now, **access** has become more of a concern, a term that took off in the 1980s.

Most of the time, the information we want is in journal **articles,** a topic that really started to grow when more of our journals became electronic. But all journal articles are not equal. We now want—no, demand—articles that show us **evidence.** The term evidence-based was the most common hyphenated word in the corpus.

Teaching has always been one of our functions. With the information world more complicated now, we are doing more **teaching, training,** and **instructing.** This is both to ourselves and to our users. The trend line for these combined terms reveals an almost straight-line increase in our use of teaching words. I predict that teaching will take on an even more important role for our profession.

Moving now to the biomedical environment outside our doors: even though **medical** is the first part of our association's name, there has been decreasing use of the words medical and **medicine** since 1961, as we have expanded our horizons and exhibited a steadily growing interest in **health** and **healthcare.** These latter terms were barely mentioned in 1961. I think this growth represents a great awakening of our profession. We are not just concerned with our little bibliographic world any more—we are concerned about what our profession and our skills can do to increase health in the world. Indeed, not only do many of us now call ourselves *health sciences* librarians, our theme for next

year's annual meeting is ''One Health: Information in an Interdependent World.'' Figure 4 is an overlay of the concepts medical and health. We can see that their trend lines crossed in the 1990s. This is a significant change in our vocabulary and was one of the more fascinating discoveries for me.

As we became more interested in health, we began focusing on the **clinician** and things **clinical** in nature. There was very little use of these words in the early 1960s. To better work with clinicians, we are literally stepping outside our building now, joining them on their own turf. For a while, we employed the concept of clinical medical librarians (**CMLs**), but the term **informationist** is now on the rise.

More evidence of our interest in and outreach beyond our traditional audience can be seen in two other groups. We have a steadily rising interest in both **consumers** of health information and in what I call our ultimate customer—the **patient.** Interest in these groups was practically nonexistent to our predecessors fifty years ago.

Management

This category surprised me the most. I was not surprised to see the other three categories form themselves from the sparklines, but it was only after the word frequency analysis that I realized how many management terms had entered our story. We do not write much about management per se, but increasingly, we are using these terms to tell our story from a specific management point of view. Yes, we have always had to manage our libraries, but we did not speak much of managing in the 1960s. Our use of terms such as **manage, managing,** and **management** peaked in the mid-1980s and has since leveled off.

As I examined the rest of the words in this category, I realized that many of them experienced significant growth in the 1980s and 1990s. Why was this? This was the time when bestselling management books entered the popular culture. You could not attend a library meeting in the late 1980s without somebody mentioning *The One Minute Manager, The 7 Habits of Highly Effective People, Who Moved My Cheese?,* or *The Essential Drucker.* Our profession took to these, and it changed our vocabulary. Rather than analyzing these words individually, I have woven them into a paragraph to illustrate our new management vocabulary. While now commonplace, these words were barely used in the 1960s and would have sounded quite odd to our ancestors:

> We are now faced with ever more **challenges** and **barriers** in our environment. But perhaps these just represent more **opportunities**. We can be successful by **collaborating** with **partners** or by using a **team** approach. If we can just stay **focused** on our **mission** and implement our **strategies**, we can achieve our **goals,** hit our **targets,** and **enhance** our **outcomes.**

We are no longer running our libraries like academic departments. We are running them more like businesses, and our changed vocabulary shows this.

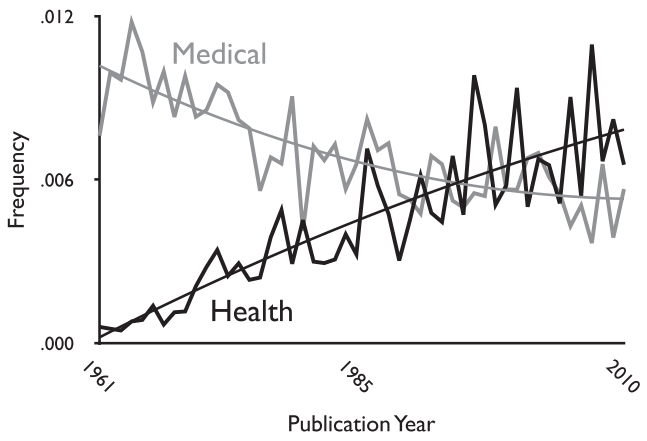
Technology

A single root word popped out of my analysis that I think dramatically demonstrates the changes that technology has played in health sciences libraries. That root is **alphabet,** which includes alphabetical and alphabetizing. As computers automated many of our processes, having to alphabetize things has became such a rare job that we barely speak of it now.

Technology has affected us personally as well as professionally. I want to get a quick visual. If you have an electronic device on you, could you turn on the screen and hold it up? [The audience raised hundreds of tablets, laptops, and smart phones.] Behold your fellow geeks. We are definitely tool users. As you might expect, there are a lot of rise-and-fall graphs in this category. We quickly pick up new tools to try them out. Then, we write about how we use them in our libraries. That is the rise. In many cases, we drop a technology when newer better technology replaces it, and that is the fall. For example, we rarely hear of people sending telegrams anymore, because that technology was replaced. However, in other cases, older technology is still being used, but it has become so embedded into our lives that we do not even bother to mention it any more. We just assume it is there. I might say, ''I spoke with Gabriel Rios last week,'' and not bother to mention that this was on the telephone. It is just assumed, since I live in New York and Gabe Rios lives in Alabama. These two scenarios—replaced technology and assumed technology—produce similar rise-and-fall graphs but are based on two different circumstances.

We obviously talk more about **technology** now than we used to. Although it is true that computers are what really changed our world, it was surprising to find out how steady our use of the word **computer** has been over the decades. Computers are nothing new to us. In earlier times, we tended to talk about

Figure 4
Change in usage frequency* between words for medical (medical, medicine, etc.) and words for health (health and healthy)



* The number of times the words appeared in a given year divided by the total number of words for that year.

bringing **machines** into the library to make our tasks easier, but that terminology has mostly died out.

In the 1970s and 1980s, we were heavily invested into **automating** various activities in our libraries. In most of our libraries, these processes are now automated, so we do not write about them as much. That is the embedded factor at work. Now, we are far more interested in talking about things **digital** and **digitizing** things. Figure 5 is an overlay showing how our interest has shifted over the years from automation to digital.

In the early days of automating our libraries, we were very interested in the **hardware** involved. We do not use the term hardware as much nowadays. Of course, you cannot do anything with hardware without **software,** and our interest in software increased in the 1980s and 1990s. It was fascinating to see how hardware and software were of equal interest to us in the 1960s and 1970s, but we soon became far more interested in the software. It is not just the shiny objects that attract us—it is what the software allows us to do with those shiny objects. Interest in both of these terms has decreased, as more of our activities have moved to the **Internet**, which exploded onto our professional scene in 1993. After that initial explosion, our use of the word has dropped in recent years. Does that mean we are no longer interested in the Internet? No, it means we have shifted our vocabulary to now concentrate on the **web** and **web-based** information. The overlay in Figure 6 reveals our shift in vocabulary usage from Internet to web.

Shifting to a web-based environment brought us new concerns. We cannot just point to the stacks to get people to their information anymore. We have to be concerned about our users **navigating** around the web. So we focus on improving **interfaces** to make navigation easier. It was interesting to see that we have been talking about interfaces even before the web came around. We have always been interested in making things easier for people to use, which is why **usability** is a term that has recently come into vogue.

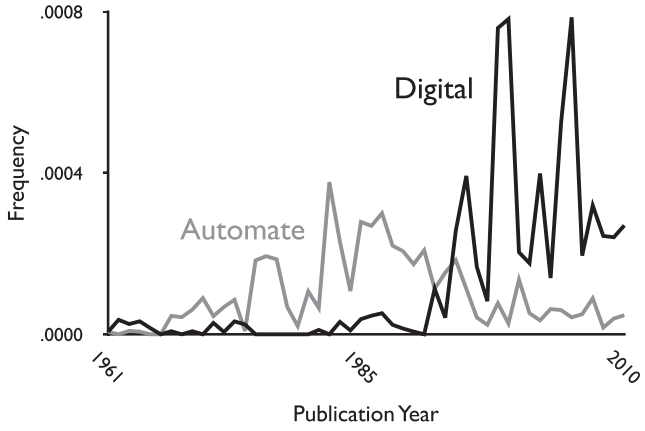
I'm going to step back from the Internet and talk about our office technology for a while. **Photocopiers** were new technology in the 1960s, and we talked about using them in various processes, primarily interlibrary loan. Now, photocopying is such a common thing that we rarely mention it anymore. **Fax** gathered a small amount of interest in the 1960s and 1970s, but it really took off in the late 1980s. Now, with scanning and PDFs commonplace, we do not talk much about faxing anymore.

A ubiquitous, if not addictive, tool for all of us today is **email.** It is hard to recall how things were accomplished without it. While our use of email is still growing, at least as far as my inbox is concerned, it looks like our writing about it has leveled off, and I have to think this will eventually fall, as it becomes an embedded technology.

In my word analysis, anything that had a space or a punctuation mark behind it was counted as a word. So I have lots of numbers and symbols counted as words, most of which appeared just a few times each. But there was one particular symbol that appears almost 5,000 times in the corpus: the **asterisk.** Up until 1989, it seldom appeared. But after that its usage exploded and shows little sign of decreasing. Strangely enough, this is another technology effect. Can anyone name what popular software came onto the market in 1987? It was PowerPoint, and the asterisk is the OCR translation of a bullet point. PowerPoint hit our profession hard, enough for it to make a substantial effect on how we communicate, even in our printed journal.

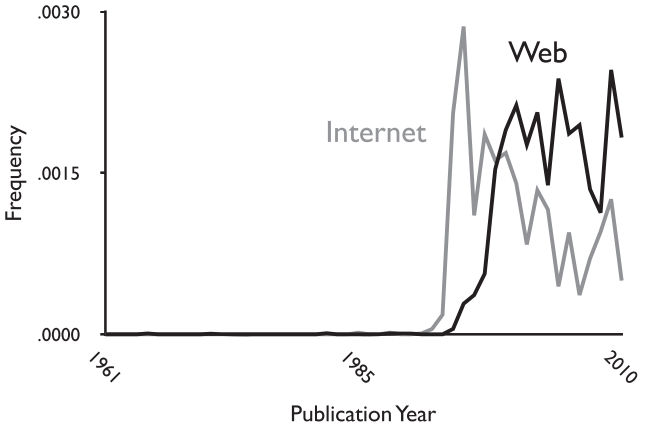
I have talked quite a bit about computer technology, but there was another type of technology in the 1970s that many thought might eventually replace books. This was the world of **audiovisuals** (**AV**). **AV** came on strong in 1969, but peaked in 1984, and by

Figure 5
Change in usage frequency* between words for automate (automate, automation, etc.) and words for digital (digital, digitize, etc.)



* The number of times the words appeared in a given year divided by the total number of words for that year.

Figure 6
Change in usage frequency* between the word Internet and words for web (web, website, etc.)



* The number of times the words appeared in a given year divided by the total number of words for that year.

the late 1990s, it was barely mentioned. In the 1960s, **film** was thought to be the new wave of delivering health information, but it quickly dropped out of sight by the 1980s. It was just too difficult to discover, acquire, and make available to users. **Slides** were much cheaper to produce than film, but except for a brief run in the 1970s, they too faded. **Videotapes,** followed by the very hi-tech **videodiscs,** were poised to become the next big thing in AV, but except for a huge blip in 1991, which came from a single-article survey of videodisc trends in health sciences libraries, these too became but minor players in our story. Solving the storage, not the delivery problem, was the aim of **microfilm** and **microfiche.** But as the promise of electronic information came to fruition, our interest in these formats became micro as well.

I want to finish this technology section with the presentation of the prestigious Burj Khalifa award, for the technology that had the sharpest rise and fall in usage. This award is named after the tallest building in the world, the Burj Khalifa in Dubai, and the winner's graph will most closely match the silhouette of the building. We have some great nominees today, all of which enjoyed a brief time in the spotlight. They are: CD-ROMs, those shiny discs we used to love; Mosaic, the first publicly available browser that existed before Netscape and Internet Explorer; personal digital assistants (PDAs), now all sitting in drawers somewhere; and Gopher, the Internet-based protocol for distributing documents, developed by the University of Minnesota. While the judging is already complete, I thought you might like to try guessing which technology actually gets the award. I will use Poll Everywhere, which lets you vote by texting, so please vote now. It appears that the audience has chosen PDAs as their guess to win the Burj Khalifa award, so let us look at the nominees to see what the judges said.

CD-ROMs had a terrifically sharp rise, but their fall was too gradual to fit the silhouette.

Mosaic had a very Burj Khalifa–like shape, but the judges took off points for both the low number of times the word was used and for that alternative meaning bump in the early 1960s.

PDAs also had a very similar silhouette to the Burj Khalifa, but the judges deducted points for those last-gasp mentions in an article in 2010. PDAs just did not stick the landing.

Which means that the winner is **Gopher.** [Katherine Chew accepted the award for the University of Minnesota. Congratulations, Katherine.] The judges were very impressed with Gopher, which achieved an almost perfect rise-and-fall fit with the Burj Khalifa silhouette. The University of Minnesota truly deserves this award.

Research

Before I begin talking about research, I need to make a confession. I purposefully left out part of the Estelle Brodman quote I used earlier. She began her statement with this: ''I believe a professional, quasi-scholarly journal, such as the *Bulletin of the Medical Library Association,* has several responsibilities'' [3]. While this classification was undoubtedly true in 1957, today's *JMLA* could never be considered a ''quasi-scholarly journal.''

We are publishing many more research-oriented articles than we used to. Two studies that analyzed articles, not words, substantiate this rise. Alexandra Dimitroff analyzed articles published in the *BMLA* between 1966 and 1990 [6]. Using the same methodology, Sally Gore and her colleagues did a follow-up study for articles published from 1991 to 2007 [7]. Dimitroff found that 29.8% of those earlier articles were research oriented, and Gore found that research-oriented articles had risen in the later years to 51.0%, an impressive increase.

My analysis confirms not only a steady increase in research-related words, but reveals what I'm calling the ''IMRaDification'' of our profession since the late 1980s. IMRaD stands for the basic sections of a scientific research paper: Introduction, Methods, Results, and Discussion. Look at the increase in these words, and note their pattern, especially when they started to increase. **Introduction** was fairly level until a rise in the late 1980s. **Methods** and **methodology** were also fairly level until a rise in the late 1980s. **Results** had an initial rise in the late 1960s, followed by a leveling, and another rise in the late 1980s. **Discussion** shows a similar pattern: level until a rise in the late 1980s.

While there are undoubtedly multiple factors behind this pattern, I hypothesize there is one primary reason for this increase, and that is the MLA strategic plan, which was released in 1987. Goal 3 of that plan stated, ''MLA is dedicated to improving health through professional excellence and leadership in research in health information science'' [8]. The profession has responded extremely well to that goal.

I will not do a detailed analysis of this topic. I just want to show you some other words that will further demonstrate our growth in research. Our use of the word **research** has grown considerably since the 1980s. This is good, because we seem to have many more **questions** now than we used to have, and we need answers.

We follow the typical path of research studies. Each study first requires a **review** of the literature and an examination of past relevant **studies.** A methodology must be decided upon. **Surveys** have been growing in popularity over the years, although there has been recent growing interest in **bibliometric** studies. **Participants** in the research are identified. Increasingly, the studies have employed **randomization** of some kind. With luck, the results are statistically **significant,** and we can draw a valid **conclusion** from the study.

FUTURE DIRECTIONS?

When I first started analyzing the sparklines, I thought I would present them by their archetypal shape, but I soon realized that the four categories I just went through told a more coherent story. I did, however, want to show you a subset of the rise archetype, consisting of what I call hockey stick terms

because the shape of their graphs resembles a hockey stick. These are words that had almost no usage, until a sharp uptake in recent years. These words *may* indicate future trends for our profession, but the problem with these words is that we do not know whether they will continue their rise in usage or become a rise-and-fall word, perhaps becoming eligible for a future Burj Khalifa award. Take a look at these hockey stick terms, and you will get a chance to vote on the one you think most likely to succeed.

Bioinformatics and **informatics** have been growing, with a huge spike in 2006, but that was followed with a sharp drop. Are they part of our future? **Blogs** are hot now, but recent evidence shows a cooling off in librarians' interest. **Cochrane** certainly fits with our desire for evidence-based literature. Will this interest continue? Electronic health records (**EHRs**) are brand new to our story, and they are white hot now. Will they stay hot? **Google:** love it or hate it, we are talking more about it now. Will it become so embedded in our lives that we will not talk about it anymore? **Literacy** is a triple threat. We talk about health literacy, information literacy, and computer literacy. This could stick around, although it had a surprisingly sharp drop in 2010. Does **sustainable** have **sustainability?** This is a hot management term that could have legs, or it could be discarded as just another buzzword.

Using live polling, the audience has chosen EHRs as the term most likely to succeed, followed by bioinformatics. At the end of my talk, I will have a sign-up sheet. I am looking for a volunteer to come to our annual meeting in 2062 and report on how these terms did over the next fifty years.

SUMMARY

Why do we study history? We live in the present, and we plan for the future, so what good is the past? History lets us understand change and how our current situation came to be. It answers that basic question—to paraphrase the Talking Heads—''You may ask yourself, well, how did we get here?''

Our words, as printed in the *BMLA* and *JMLA*, were carefully chosen by hundreds of individual authors to tell their particular stories. But in the aggregate, their words reveal the story of our profession as it has changed over the years. Let me summarize some of the major differences between ''here'' and fifty years ago.

- We have seen a massive de-emphasis on the physical objects that used to be our stock in trade.
- Instead of those packages, we are much more concerned with information, which increasingly is digital in format.
- We understand that not all information is created equal. We prefer our information to be evidence based.
- We have a much larger worldview, with a growing emphasis on health, not just medicine.
- With this larger worldview, we serve a much larger audience, and we are now reaching out to clinicians, consumers, and patients.
- To better serve that audience, we are leaving our physical building, interacting with our users in their environment.
- Instead of just giving people information, we are much more involved in teaching them how to find and evaluate information on their own.
- We are running our libraries more like businesses, using concepts pioneered in the management field.
- We love technology. We quickly pick up new technology to see if it will be useful to us. We also quickly drop it when better technology appears.
- We are no longer satisfied with ''how I do it good'' reports from librarians. We publish more research papers using the standard IMRaD structure of scientific journals.

While history can explain how we got here, it does not do a great job of telling us what happens next. It can hint at but cannot reveal the future. Warren Buffet said, ''If past history was all there was to the game, the richest people would be librarians'' [9]. We know that is not true. Indeed, there is much more to the game than past history. We are rightfully concerned about the future.

However, as the great baseball philosopher Yogi Berra supposedly said, ''It's tough to make predictions, especially about the future.'' While our story is rich, with lots of plot twists and character development, a basic existential question remains for us: Is ours a never-ending story? Since it is being written every day, we cannot skip chapters to see what happens next. Some believe that the semantic web, combined with massive cost cutting in health institutions, will lead to our profession's decline and possible demise. Others believe that the increasingly complex information space of biomedicine will not only require our skills, but also lead us into a great age of librarians.

What do I believe? I believe…I believe I'll have another cookie.

ACKNOWLEDGMENT

The author thanks Octavio Morales, head of computer services at the Samuel J. Wood Library, for his essential work in creating the PERL script that counted and sorted the words from fifty years of text files and for creating the interactive website displayed during the lecture.

REFERENCES

1. Kronick DA. The librarian's life, scholarship and librarianship. Bull Med Lib Assoc. 1980 Oct;68(4:):327–35.
2. Michel JB, Shen YK, Aiden AP, Veres A, Gray MK, The Google Books Team, Pickett JP, Hoiberg D, Clancy D, Norvig P, Orwant J, Pinker S, Nowak MA, Aiden EL. Quantitative analysis of culture using millions of digitized books. Science. 2011 Jan 14;331(6014):176–82. DOI: http://dx.doi.org/10.1126/science.1199644.
3. Brodman E. Hail and farewell. Bull Med Lib Assoc. 1957 Jul;45(3):426–8.
4. Tufte ER. Beautiful evidence. Cheshire, CT: Graphics Press; 2006.

5. Plutchak TS. Breaking the barriers of time and space: the dawning of the great age of librarians. J Med Lib Assoc. 2012 Jan;100(1):10–9. DOI: http://dx.doi.org/10.3163/1536-5050.100.1.004.
6. Dimitroff A. Research in health sciences library and information science: a quantitative analysis. Bull Med Lib Assoc. 1992 Oct;80(4):340–6.
7. Gore SA, Nordberg JM, Palmer LA, Piorun ME. Trends in health sciences library and information science research: an analysis of research publications in the Bulletin of the Medical Library Association and Journal of the Medical Library Association from 1991 to 2007. J Med Lib Assoc. 2009 Jul;97(3):203–11. DOI: http://dx.doi.org/10.3163/1536-5050.97.3.009.
8. Medical Library Association. Shaping the future: the strategic plan of the Medical Library Association. Chicago, IL: The Association; 1987.
9. Buffet M, Clark D. The tao of Warren Buffett. New York, NY: Scribner; 2006.

AUTHOR'S AFFILIATION

Mark E. Funk, MA, AHIP, FMLA, mefunk@med.cornell.edu, Associate Director, Resources Management, Samuel J. Wood Library, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10065